「資料の活用」領域での教材研究
—— 中学一学年 数学での新学習指導要領への対応 ——

A research on teaching the domain "Practical use of data"
—— corresponding to the New Course of Study on mathematics
for first grade students of junior high school ——

木 下 淳 子　　門 田 良 信
KINOSHITA Junko　　KADOTA Yoshinobu
（海南市立下津第一中学校）（和歌山大学教育学部数学教室）

（2009 年 10 月 5 日 受理）

新学習指導要領の中学校数学では，従来の領域に「資料の活用」が加わった。平成 24 年度の実施に向けて本年度（平成 21 年度）から移行措置が始まり，各教科書会社から 1 学年の生徒全員に「補助教材」が配布された。この補助教材を手がかりに，資料の活用について授業を行うときに出会う問題点を考える。この領域では様々な現象を直接扱う方法が主題となっていて，他の領域の数学とは異なる要素を含んでいる。教材の選択，考え方や感じ方の違い，コンピュータの活用など授業を展開する上で多様な配慮や注意が必要であることが分かった。

キーワード: 新学習指導要領, 中学校数学, 資料の活用, 度数分布表, 代表値, 有効桁, 10^n

1. はじめに

平成 24 年度に実施される中学校の新学習指導要領 [4] の数学では, 従来の領域「数量関係」が 2 つに分かれて, 新領域は「数と式」,「図形」,「関数」,「資料の活用」の 4 つとなった。

教育目標では「表現する能力」や「数学的活動」が重視された。特に後者については, 各学年ごとに次のように述べられている。

「(上記 4 領域）の学習やそれらを相互に関連付けた学習において, 次のような数学的活動に取り組む機会を設けるものとする。

ア. 既習の数学を基にして, 数や図形の性質などを見いだす活動

イ. 日常生活で数学を利用する活動

ウ. 数学的な表現を用いて, 自分なりに説明し伝え合う活動」

上記は 1 学年の活動について述べたもので, 2, 3 学年では部分的に異なる。（イ）では（日常生活）が（日常生活や社会）となり, （ウ）では（自分なりに）が（根拠を明らかにし筋道を立てて）と言い換えられている。

「資料の活用」の内容は, 1 学年については「 目的に応じて資料を収集して整理し, ヒストグラムや代表値の必要性と意味を理解して, 資料の傾向をとらえ説明できること」となっている。

2 学年の内容は「確率」, 3 学年では「標本調査」と続く。「確率」は従来から指導してきたものであり, 1, 3 学年の内容との関連が求められると思われる。「標本調査」の内容については現時点では不明な部分が多い。

現行以前の学習指導要領では「確率」の他に, 資料の整理, 相関図・相関表, 標本調査に関する学習が行われていた。今回の改訂がそれと異なるのは, 数学的活動の（イ）や（ウ）に関連して提示されたことにある。「活用」という言葉はその意味で使われていると思われる。

平成 24 年度に実施される中学校の新学習指導要領に向けて，今年（平成 21 年度）から移行措置が始まった。1 学年の生徒全員に各教科書会社による「補助教材」が配布されている。

また, 小学校の新学習指導要領 [6] でも「資料の平均や散らばりを調べたり, 度数分布を表す表やグラフを用いて表現したり, 統計的に考察したりする」ことになった。こちらは 23 年度から実施される予定であり, その内容は現時点で明らかではない。

上述した状況の中で, ここでは補助教材を手がかりに「資料の活用」の 1 学年の内容に焦点を絞って, 授業を展開するときに出会う問題点を指摘し考察する。

次の第 2 節では, この領域のもつ他の領域の数学とは異なる特徴をまとめ, ここでの問題意識を明確にする。3 節から 6 節までは, 各単元に現れる概念の意味とそれらを求める方法について考察する。第 3 節では, 度数分布やヒストグラムの作成について, 第 4 節では, 代表値 (平均値, 中央値, 最頻値) や範囲について, 第 5 節では, 統計調査の大まかな計画について, 第 6 節では, 近似値について考える。第 7 節では自由研究として, 桜の開花時の和歌山市と札幌市の一日ごとの最高気温を比較する問題を与える。第 8 節では, 第 1 節から 7 節で述べたいくつかの事項に関連して授業の準備として知っておきたいことを述べる。

2. 考察のポイント

前節では「資料の活用」の意味を考え，現時点での状況を見た。そのことから気づくことは次の 2 点である。

(1) すべての資料が出揃っているわけではないが, 今年から授業を行わなければならない。

小学校での移行措置は当面期待できないし, その後も小学校との円滑な接続のために努力しなければならない。他の領域との関係や領域としてのまとまりにも気をつけたい。

(2) 社会生活を送る上では様々な数値, 度数分布, グラフなどを目にする。ともすれば現代社会では様々な数字が一人歩きをして, 元になる資料の信頼性が忘れられる傾向もある。それらの情報の中には教材ととして面白いものも含まれているはずであるが適切なものは簡単には見つからない。

このことに関連して, 補助教材や教科書の段階から資料が与えられた時には出典をチェックする習慣をつけておきたい。資料の信頼性は最も重要であり，「作成者, 作成時期, 調査対象, 調査場所, 方法」などを納得しておく必要がある。また, 生徒が観察や調査を行って資料を採取して開示する場合には, 必ずそれを記入するように指導すべきである。

さらにアンケートなどの場合には, 個人情報の保全に注意し, 集めた情報は, 協力者が納得した目的以外のことで使ってはならないことは, 常識として生徒に注意しておくべきことである。

このようなことは補助教材でも軽く触れられている。導入の部分で生かしていきたいと思う。

また, 社会生活上の面白く有用な問題を探そうとすれば, 教材研究は難しく, そのような問題作りには様々な教育上の配慮が必要とされる。

数学では公式を導いたり, それを使って問題を解くことが主な内容となる。しかしこの領域では様々な現象を直接扱う方法が主題となっていて, 感触が違う。内容的に異なる要素をまとめておく。

(3) 度数分布やヒストグラムを作成したり, 代表値（平均値, 中央値, 最頻値）や範囲を求める学習では, 決まった手順や同種の計算を繰り返し行う作業が多い。また, その結果は目的によって変わり得るし, 常に一意的な解答が得られるわけではない。

単純で長い計算をする場合には電卓やコンピュータを用いるべきである。補助教材の問いなどでは, 手計算でもできるように配慮されている。しかし手計算でやることによって, この学習を「面倒くさい」,「面白くない」という印象を生徒に与えないようにしたい。

この単元に現れる様々な方法はすべての資料に対して使えるものである。ただし, それらが常に有用な結果を導くとは限らない。資料を採取した目的によって, 最頻値に注目すべきか平均値に注目すべきかが異なる。また, ヒストグラムと度数多角形のどちらが見やすいのかも場合によって異なる。そのような事情から, それらを求める方法と並んで一つ一つの概念の意味するものを深く理解しておく必要がある。

(4) 「活用」ということでは, (3) の結果を得た上でそれのもつ意味を検討する必要がある。一つの問題が長く続き, その理解や解析に様々な労力を要する。さらに, 根拠を明らかにしつつ判断したり説明したりする能力が要求されている。

ここではいわゆる数学的能力の他に, 様々な問題に対応できる柔軟な思考力や適切な判断力, 国語的表現力なども必要とされる。

補助教材では，導入の部分で資料と大まかな問題を提起し最後の自由研究や発展問題などで解くような工夫が施されている。

上記の (1) と, (2) の教材研究に関すること以外は授業内容と直接的な関係が薄いから, 以下では取り扱っていない。(3) の計算機の導入については, 工夫を要する事項でありここで一括して扱える問題ではない。

以下の節では, (3) の後半部分の様々な概念の意味することと (4) の教材研究に関することを中心として, 授業において気をつけたい問題点を考えていく。

3. 度数分布表とヒストグラム

最初にここで使う用語についての解釈を与えておこう。

資料 : 観察, 実験, 調査などによって得られる数値のまとまり
測定値 : 資料を構成している個々の数値
分布 : 観測値の全体としての大きさと散らばり具合
度数 : 一定の性質を満たす測定値の個数
階級 : 測定値の大きさによって資料を分割するときの一つの区分

3.1 度数分布表

(1) 資料の全体的傾向や特徴を理解するための最も良い方法の一つは度数分布表を作成することである。度数分布表は, 測定値を大きさによって分類したものだから, これを作成する意義は次のものであると考える。

(i) 測定値の大きさや散らばり具合を見ることができる。
(ii) 平均値, 中央値, 最頻値の大体の値が分かる。また, 平均値を計算する労力を軽減できる。

(2) 度数分布表の作成

例 **3.1** 次の資料はのぞみさんのクラス 28 人が 1 学期に図書館で借りた本の冊数を調べたものです。度数分布表を作りなさい。

表 **3.1**

6	3	9	4	10	9	10	11	7	6
12	5	11	9	8	10	2	7	8	9
6	5	9	15	3	9	12	9		

(i) 資料の最小値と最大値を求める。
(この例では 最小値 2, 最大値 $b = 15$ となる。)

(ii) 階級の幅を適当にとって (i) で求めた最小値に次々に加えて, 境界値を求める。
（ここでは階級の幅を 3 にとることにする。境界値は 2, 5, 8, 11, 14, 17 となる。）

(iii) 階級値は境界値に階級の幅の $\frac{1}{2}$ を加えて定める。
（階級値は 3.5, 6.5, 9.5, 12.5, 15.5 となる。）

(iv) 各階級に属する度数を資料に基づいて調べる。測定値が境界値に等しいときには, 大きい方の階級に入れる。
（度数は 4, 7, 12, 4, 1 となる。）

(v) 度数の和が測定値の総数に等しいことを確かめて「合計」の欄に記入する。
$(4+7+12+4+1=28)$

表 **3.2**

階級値	境界値	度数調べ	度数
3.5	$2_{\text{以上}}$ ～$5_{\text{未満}}$	////	4
6.5	5 ～8	卌 //	7
9.5	8 ～11	卌 卌 //	12
12.5	11 ～14	////	4
15.5	14 ～17	/	1
合計			28

(3) [注意] (i) 上記 (v) において度数の和が測定値の総数になることを確かめるのは, (iv) の 度数の数え方が正しいことを確認するためである。

(ii) ここで示したものは, 小さい方の境界値以上大きい方の境界値未満の度数分布表を作成する手順である。補助教材では常にこの方法で作成される。もう少し一般的な度数分布表を作成する方法を第 8 節に記す。

(iii) 度数分布表の手順 (iv) のために, 一回くらいは正の字あるいは 卌 の欄をつけた度数分布表を提示しておいた方がよい。
（第 2 節の (3) でコンピュータを使うことと矛盾しているように見えるが, 度数を求める手間には独特なものがあり, 一度は自分で経験しておくべきであると思う。）

(iv) 測定値を小さい順に並べ変えた表を作っておくと, 度数分布表の作成, 後に出てくる平均値, 中央値, 最頻値の導出にも便利である。

(v) 度数分布表の形は [5] で指摘されているように, 境界値のとり方や階級の分け方によって様々に変わり得る。そのことも重要である。

3.2 ヒストグラム

ヒストグラムは度数分布表に現れた資料の特徴を視覚的に把握できるので, 実感を伴ってその特徴を理解できる。

ヒストグラムの横軸には境界値を記入する。（第 8 節参照。）

3.3 度数分布多角形

(4) 度数分布多角形は現行以前の学習指導要領の中にあったので, どの補助教材でも取り上げている。ヒストグラムは棒グラフで表したものであるが, 度数分布多角形は度数を折れ線で結んだものである。これを作成する意義は次のように考えられる。

(i) 2 つのヒストグラムを比べるときに便利。

(ii) 全体の変化の様子やゆがみ, とがり具合が直観的に分かる。（ となるか となるか。）

(5) [注意] (i) 度数分布多角形は折れ線グラフではないから縦軸の目盛りは 0 から書き始める。

(ii) 度数分布多角形は横軸上の点から始まり横軸上の点で終わる。したがって, 度数 0 の階級を両端に付け加えた形で描く。

3.4 相対度数

(6) 相対度数分布表を作成する意義は次のものであると考える。

(i) 資料の中で各階級値をとる度数の割合が分かる。

(ii) 測定値の総数の異なる 2 つ以上の資料を比べることができる。

(7) [注意] (i) 相対度数分布表の「合計」は一般には 1.00 にはならない。次のように対応する。

(a) 相対度数分布表の「合計」の欄にはそのままの値（例えば 0.99 や 1.01 ）を記入し, 表の下に （合計が 1.00 にはならないのは四捨五入による計算誤差）と記す。

(b) 最大相対度数のところで調節して必ず 1.00 になるように書く。

(ii) 相対度数分布表は離散確率分布と関係をもつ。（第 8 節参照。）

4. 代表値と散らばり

資料のすべての測定値の大きさを代表させて一つの数値を用いて表したものをその資料の代表値とよぶ。代表値は何らかの意味で資料の中心となっていて, 様々なものが考案されている。

この節では資料または度数分布表が与えられているとする。

4.1 平均値

(1) 資料に対して平均値（または平均）は次の式で計算される。

$$(\text{平均値})=\frac{\text{測定値のすべての和}}{(\text{測定値の総数})} \tag{4.1}$$

度数分布表に対して平均値は次の式で計算される。

$$(\text{平均値}) \\ =\frac{\{(\text{階級値})\times(\text{度数})\}\text{ の階級に関する和}}{(\text{測定値の総数})} \tag{4.2}$$

(2) 直観的な意味で「平均値は測定値をならした値である」といえる。このことを次の例 4.1, 例 4.2 で説明しよう。ただし, 例 4.1 の考え方は現行の小学校で履

修されている。

例 **4.1** ((4.1) の場合) 資料 3, 6, 5, 2, 7, 4 の平均を求めなさい。

平均は (4.1) より

$$\frac{1}{6}(3+6+5+2+7+4)=4.5$$

となる。

平均値の意味を考えるために, 例 4.1 の測定値の棒グラフを作る。その中に水平な線を引くと, 両端の棒の線またはその延長と交わる長方形ができる。この水平線を基準として棒グラフのはみ出した面積を足りない部分に埋めてやると, ちょうど過不足なく長方形の面積ができる水平線が引ける。水平線の高さ（ここでは 4.5）が平均値となる。これより, すべての測定値の棒の長さが平均値によってならされたことになる。

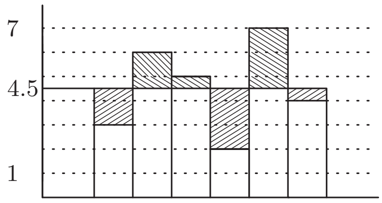


図 **4.1**

例 **4.2**((4.2) の場合) 度数分布表 3.2 の平均を求めなさい。

表 3.1 を数値の小さい順に並べ替えると次のようになる。

表 **4.1**

2	3	3	4	5	5	6	6	6	7
7	8	8	9	9	9	9	9	9	9
10	10	10	11	11	12	12	15		

表 4.1 の各数値をそれに最も近い階級値に置き換えると次のようになる。

表 **4.2**

3.5	3.5	3.5	3.5	6.5	6.5	6.5
6.5	6.5	6.5	6.5	9.5	9.5	9.5
9.5	9.5	9.5	9.5	9.5	9.5	9.5
9.5	9.5	12.5	12.5	12.5	12.5	15.5

表 4.2 では 3.5 が 4 回, 6.5 が 7 回, 9.5 が 12 回, 12.5 が 4 回, 15.5 が 1 回現れている。これが表 3.2 の度数分布表となっている。これより表 4.2 の置き換えによる誤差を無視すれば次の式が成り立つ。

((4.1) による表 4.2 の平均)

$$=\frac{1}{28}(3.5+3.5+\cdots+12.5+15.5)$$

$$=\frac{1}{28}(3.5\times 4+6.5\times 7+\cdots+15.5\times 1)$$

=((4.2) による表 4.2 の平均)

例 4.1 で図 4.1 を作ったように表 4.2 から図 4.1 に相当する図を作ると, 水平線の高さが平均値となる。これより (4.2) は度数分布表の階級値を度数についてならしたものとなる。

(3) [注意] (i) 資料から計算した平均値と度数分布表から計算した平均値は, 似てはいるが一般に異なる値となる。(この場合, 表 4.1(表 3.1) の平均値は 8.0, 度数分布表 3.2 の平均値は 8.54 。)

(ii) 一般に測定値がすべて整数値でも平均は小数となる。

(4) 例 4.1 の考え方から「平均の平均は平均」とでも言いたくなる次の問題ができる。これも小学校で習っているはずであるが, 確率の場面ではこのような考え方が大変重要となるので敢えて書いておこう。

問 **4.1** さいころを 10 回振ったときの平均は 3.1 でした。そのあと 5 回振ったときの平均は 3.4 でした。この実験を 15 回続けて振ったと考えたときの平均を求めなさい。

(5) 天びんの棒の両端に物をのせるとき, 支点において釣り合う。支点の位置が両端にのせた物の平均値を表すことになる。例 4.1 と同様の数値を使って次の問題ができる。

問 **4.2** 等間隔に目盛りのついた棒がある。棒の一方の端からそれぞれ 30, 60, 50, 20, 70, 40 cm の位置に同じ重さのおもりをつるす。このときの支点の位置を求めなさい。ただし, 棒の重さは考えないものとする。

度数分布表から平均値を求める場合にも天びんの考え方は使える。今度は, 度数がおもりの重さ表し, 階級値がそれを置く位置を与える。度数分布表 3.2 と同じ数値を使って次の問題ができる。

問 **4.3** 等間隔に目盛りのついた棒の一方の端から 35, 65, 95, 125, 155 cm の位置に, それぞれ 4, 7, 12, 4, 1 グラムのおもりをつるす。このときの支点の位置を求めなさい。ただし, 棒の重さは考えないものとする。

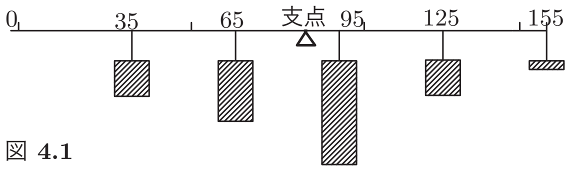


図 **4.1**

4.2 中央値 (Median)

中央値は次のようなものである。

(i) 測定値を小さい順に並べた真ん中の値。真ん中に 1 つの測定値がない場合には, 真ん中にある 2 つの測定値の平均値。

(ii) ヒストグラムの両側の面積が等しくなるように縦線を引くと, その線の横軸上の値が中央値となる。

4.3 最頻値 (Mode)

最頻値は資料の中で最も多数回出てきた測定値である。

(i) 資料を一べつしたときに最も目に留まる値。

(ii) 補助教材 [1, 2, 3] の中にはないが, 実際には最頻値が 2 つ以上ある場合もある。

4.4 範囲 (**Range**)– 散らばり–

範囲は資料の代表値ではなく散らばり具合を表す統計量の一つである。

(i) 範囲 = 資料の最大値 – 資料の最小値

(ii) 「範囲」は大体の「散らばり」を表す 1 つの数値に過ぎない。かけ離れた数値が一つでもあれば「範囲」は大きくなるが, 直観的にはより小さな散らばりをもつ資料もありうる。

4.5 代表値の選び方

平均値, 中央値, 最頻値の中で統計で最もよく使われるのは平均値である。しかし, それらには代表値としてそれぞれの特徴があるので目的によって使い分けられるべきである。

(6) 平均値は代表値としてよく使われるが難点もある。

(i) 大きな資料では計算が面倒となる。

(ii) 飛び離れた測定値がある場合や分布に偏りがあると, 平均値がそれに引きずられて感覚的な「真ん中辺の値」とならない。(例えば, 誤った測定値がある場合。分布に偏りがある場合として, 所得に対する人数では, 少数の高所得者の影響が出て大半の人が平均値以下の所得になることはよく知られている。)

(7) 中央値は資料の最大値, 最小値と合わせると分布の偏りが分かりやすい。(例えば, 最大値 90, 最小値 10, 中央値 40 ならば測定値は小さい方に密になっている。)

(8) 最頻値には, 平均値や中央値のような資料の釣り合いの意味がない。感覚的に重視される値となる。(例えば, T シャツを工場生産するときにサイズ S, M, L の数値を決める場合, レストランの 一人前の分量を決める場合など。)

(9) 代表値の特性について理解を深めるために, PISA の問題を参考に次の問題を考えた。第 2 節の記述に関連して現実性を意識した。

問 4.4 のぞみさんのクラスで漢字のテストがあり, 0 点から 10 点までの整数によって採点されました。クラスの生徒は 22 人ですが当日は 2 人が欠席しました。翌日, 欠席していた生徒が登校したので彼らも同じテストを受けました。

最初の採点では最高点は 9 点, 最低点は 2 点でした。平均値は四捨五入すると 4.4 点でしたが翌日に受けた 2 人の分も含めると 4.3 点になりました。中央値も 4.5 点でしたが 4 点になりました。その 2 人の点数はクラスの最高点でも最低点でもありませんでした。

翌日に受けた 2 人の点数についてどのようなことが言えるでしょうか。

5. 資料の活用とコンピュータ

(1) ある目的をもって資料を収集・整理しその結果を発表するためには, 最初にプロセスの全体にわたる計画を立てておくことが重要となる。プロセスは次のようになる。

(i) 調査の目的を明確にする。

(ii) 目的に沿った綿密な調査計画を立てる。

(iii) 資料の収集をする。

(iv) 資料を整理する。

(v) 資料を基礎として考察する。

(vi) 結果を発表する。

(2) これらを実行に移す段階ではコンピュータを使うと, 便利で能率よく作業をすることができる。

(i) (iii) では, 実験や観察を繰り返し行って記録したり, そのことに関係したことを書いている記事や他の人の行った記録をインターネットなどを利用して検索したりできる。

(ii) (iv) では様々なプログラムを使って, 度数分布表やヒストグラムを作成したり, 代表値などを求めたりできる。

(iii) (vi) でも様々なプログラムを使って, 自分の調べたことや資料のもつ特徴を分かりやすく表現できる。

6. 近似値

6.1 測定値と誤差

(1) 資料に現れる測定値は人間の感覚や作業を経て得られたものであるから, 多くの場合に誤差を含む。真の値は誰にも分からないのが普通である。

$$(\text{誤差}) = (\text{測定値}) - (\text{真の値})$$

であるから測定値は真の値の近似値と考えられる。

測定値は真の値を四捨五入, 切り上げ, 切り捨てのいずれかによって求められたものとすると, 真の値の範囲（誤差の範囲）と有効数字が得られる。測定値は, 有効数字の数あるいはその位取りをそろえて求めておく必要がある。

6.2 10^n 表現

(2) 大きな値をとる近似値（位取りのために下位の位に 0 を並べた数）を簡潔に表現する方法を考えよう。

$10^1 = 10$ （10 の 1 回の積）

$10^2 = 10 \times 10 = 100$ （10 の 2 回の積）

$10^3 = 10 \times 10 \times 10 = 1000$ （10 の 3 回の積）

$10^4 = 10 \times 10 \times 10 \times 10 = 10000$ （10 の 4 回の積）

……………………………………………………

と表すことにする。$10^1, 10^2, 10^3, \cdots$ をそれぞれ十の 1 乗, 2 乗, 3 乗,$\cdots$ と読む。$n = 1, 2, \cdots$ について 10^n は 1 のあとに位取りのための 0 が n 個ついた数となる。これを使えばどんな数も $a \times 10^n$ の形に表される。このとき, a は $-10 < a < 10$ にとるのが普通である。例えば

$$48611 = 4.8611 \times 10^4 \qquad 516.24 = 5.16924 \times 10^2$$

$$3590 = 3.590 \times 10^3$$

などと表される。

(3) [注意] 補助教材によっては $a \times \dfrac{1}{10^n}$ まで説明しているものもある。

(4) a, b を任意の数, m, n を正の整数とするとき次の公式 (i)∼ (iii) が成り立つことが 10^n の意味からすぐに分かる。

(i)　$(a+b) \times 10^n = a \times 10^n + b \times 10^n$

(ii)　$10^m \times 10^n = 10^{m+n}$

(iii)　$\dfrac{10^m}{10^n} = 10^{m-n}$　　（ $m > n$ のとき ）

(iv)　$(10^m)^n = 10^{mn}$

((iv) については $(10^m)^n$ とは 10^m の n 個の積のことと考える。したがって, $(10^m)^n$ は 1 の後に 0 が m 個ついた数を n 回掛け合わせたものとなるから, (iv) の公式が成り立つ。)

問 6.1 次の数を $a \times 10^n$ の形（ $-10 < a < 10$）に表しなさい。

(i)　$1.24 \times 10^3 + 0.82 \times 10^2$

(ii)　$15.4 \times 10^2 - 3210$

(iii)　$6.14 \times 10^2 - 61 \times 10^3$

(iv)　$(12^2)^3$

(v)　$(6.8 \times 10^4) \div (0.25 \times 10^2)$

(5) 10^n の形で表された数値を読みとるために次の表を頭に入れておこう。

表 **6.1**

n 乗表現	10 進法表現	読み
10^1	10	十
10^2	100	百
10^3	1,000	千
10^4	1,0000	一万
10^5	10,0000	十万
10^6	100,0000	百万
10^7	1000,0000	千万
$10^8 = (10^4)^2$	1,0000,0000	一億
10^9	10,0000,0000	十億
10^{10}	100,0000,0000	百億
10^{11}	1000,0000,0000	千億
$10^{12} = (10^4)^3$	1,0000,0000,0000	一兆

位取りのコンマを 3 桁毎にとると英語で言う, thousand（千), million（百万), billion（十億）の表ができる。

次のような問題で $a \times 10^n$ と表現する便利さを確かめる。

問 6.2 (1) 地球から太陽までの平均距離はだいたい 149600000 km, 月から地球までの平均距離はだいたい 384000 km です。これらの数値を, a を整数部分が 1 けたの数字として $a \times 10^n$ の形で表しなさい。

(2) 地球から太陽までの平均距離は月から地球までの平均距離の何倍になるかを, 有効数字 3 けたで表しなさい。

7. 自由研究

(1) 各補助教材には実際の資料を活用して, 対象とする現象を深く理解するための教材がある。

例えば「日本の人口ピラミッド」は 2 つの補助教材に出ている。見る方の問題意識によって様々な見方のできる示唆に富んだ教材と思える。「気温の比較」や「北京オリンピックの日本女子選手の身長と体重の競技別の比較」などもある。それらを真似て次の例題を作ってみた。授業作りのためのたたき台になるように願っている。

(2) [桜が開花する頃の気温変化]

出張していたお父さんが五月に札幌から帰ってきて「今年は満開の桜を２度見たよ。円山公園の桜もきれいだったなあ。」と言ったので, とおる君は驚きました。

とおる君の住む和歌山市では桜はとっくに散っています。「桜の咲く頃の札幌市と和歌山市の気温は似ているのかなあ」と思い，調べてみることにしました。

気象庁のホームページによると, 今年（2009 年）の桜（ソメイヨシノ）の開花は和歌山市では 3 月 21 日, 札幌市では 5 月 1 日でした。とおる君は桜が開花した日までの 1 日の最高気温を 30 日間調べました。

表 **7.1**　**1** 日の最高気温の **30** 日間の記録（単位 ℃）

	和歌山市			札幌市	
	日数	月日	温度	月日	温度
前期	1	2/20	12.3	4/2	8.2
	2	21	8.3	3	13.6
	3	22	14.8	4	13.0
	4	23	15.8	5	12.4
	5	24	12.3	6	10.8
	6	25	12.9	7	8.4
	7	26	13.3	8	13.3
	8	27	9.0	9	15.2
	9	28	12.2	10	17.7
	10	3/1	13.9	11	10.0
	11	2	9.2	12	15.2
	12	3	6.2	13	21.2
	13	4	10.9	14	13.6
	14	5	13.2	15	12.8
	15	6	13.9	16	9.0
後期	16	7	11.2	17	13.8
	17	8	12.9	18	13.4
	18	9	13.5	19	12.3
	19	10	15.4	20	16.5
	20	11	11.4	21	10.5
	21	12	11.9	22	15.8
	22	13	17.4	23	9.1
	23	14	17.5	24	9.2
	24	15	12.7	25	12.1
	25	16	15.5	26	6.5
	26	17	19.9	27	8.3
	27	18	22.1	28	10.7
	28	19	22.6	29	18.0
	29	20	17.8	30	23.4
	30	21	18.0	5/1	19.2

(気象庁のホームページより抜粋。ただし, 左欄の前期, 後期などは筆者による。)

とおる君はこの観測値を 15 日ごとに区切り, 前期, 後期とよぶことにしてそれぞれの最小値, 最大値, 平均値, 中央値を求めました。

表 7.2

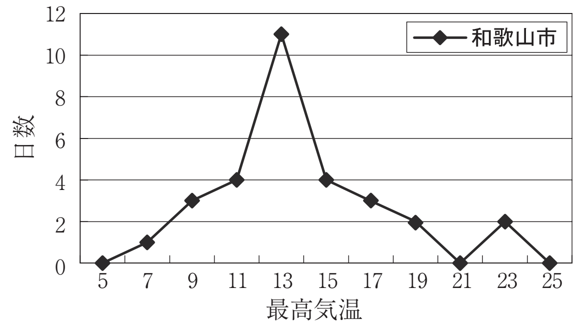
		最小値～最大値	平均値
前期	和歌山市	6.2 ～15.8	11.88
	札幌市	8.2 ～21.2	12.96
後期	和歌山市	11.2 ～22.6	15.99
	札幌市	6.5 ～23.4	13.25

2 つの市の 30 日間の最高気温の変化について考えてみよう。

(i) 表 7.2 から 2 つの市の 30 日間の最高気温について平均値と範囲を求めなさい。

(ii) 次の度数分布多角形は表 7.1 の和歌山市の最高気温を表しています。札幌市の度数分布多角形を作って, このグラフの中に記入しなさい。

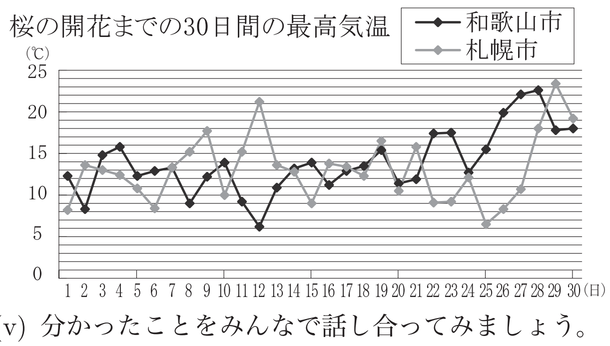
図 7.1



(iii) (i) や (ii) からどんなことが分かりますか。

(iv) 次の図 7.2 は表 7.1 を折れ線グラフにしたものです。(iii) で分かったこと以外にどんなことが分かりますか。

図 7.2



(v) 分かったことをみんなで話し合ってみましょう。

解答は様々に表現できる。解答例を第 8 節の最後に与えておく。

8. 授業のための準備

この節では前節までに述べてきたことと関連することで, 「資料の活用」を指導するときに教師として準備しておくとよいことをいくつか述べておく。授業で行う必要はないものであるから, そのための表現に関する配慮をこの節では行っていない。

8.1 境界値と階級値

(1) 第 3 節の度数分布表では, いつも下の境界値以上で上の境界値未満の表が作成される。度数分布表による表現を豊かにするためにもっと一般的な表の作成の仕方を提示しておこう。

(i) 資料の中の最小値を a, 最大値を b とし範囲 $R = b - a$ を求める。

(ii) 階級の幅 C と自然数 k を次の関係を満たすようにとる。

$$\frac{R}{k} \leqq C < \frac{R}{k-1}$$

(k または $k+1$ が階級の個数になることに注意する。 C は簡単な数にとる。)

(iii) 境界の最小値 a_0 を $a_0 \leqq a \leqq a_0 + \dfrac{C}{2}$ を満たすようにとる。

(iv) 境界値は $a_0,\ a_0 + C,\ a_0 + 2C, a_0 + 3C,$ $\cdots,\ a_0 + kC$ または $a_0 + (k+1)C$ となる。

(v) 階級値は境界値に階級の幅の $\frac{1}{2}$ を加えて $a_0 + \dfrac{C}{2},\ a_0 + C + \dfrac{C}{2}, a_0 + 2C + \dfrac{C}{2},$ $a_0 + 3C + \dfrac{C}{2}, \cdots,\ a_0 + (k-1)C + \dfrac{C}{2}$ または $a_0 + kC + \frac{C}{2}$ と定める。

(vi) 各階級に属する度数を資料に基づいて調べる。((iii) で a_0 を測定値の有効桁に端数をつけたものにとると, 境界値に等しい測定値が現われない。)

(vii) 度数の和が測定値の個数に等しいことを確かめて「合計」の欄に度数の和を記入する。

例 8.1 例 3.1 で与えた表 3.1 の度数分布表 3.2 をもう 1 度作ってみよう。

(i) 範囲は $R = 15 - 2 = 13$ となる。

(ii) 階級の数を 5 または 6 にすることに決めて $k = 5$ とすると $\dfrac{13}{5} \leqq C < \dfrac{13}{4}$ だから $2.6 \leqq C < 3.25$ を満たす簡単な数として $C = 3$ をとる。

(iii) 境界の最小値 a_0 を $a_0 \leqq 2 \leqq a_0 + \dfrac{3}{2}$ を満たす端数の付いた数として $a_0 = 1.5$ とする。

(iv) 境界値は 1.5, 4.5, 7.5, 10.5, 13.5, 16.5 となる。

(v) 階級値は境界値に 1.5 を加えて 3, 6, 9, 12, 15 となる。

(vi) 各階級の度数を資料に基づいて調べると
(階級値, 度数)= {(3, 4), (6, 7), (9, 12)
(12, 4), (15, 1)}

(vii) 合計 $(4 + 7 + 12 + 4 + 1 = 28)$

表 3.2 との違いは階級値が整数として現れ, 境界値に一致した測定値が現れないことである。平均値は 8.04 となる。

(2) 補助教材によると, ヒストグラムの横軸には境界値を記入するのが普通である。度数分布表やヒストグラ

ムを作成するまでは境界値は重要である。しかし, 平均や分散などを度数分布表から求めようとすれば，階級値の方が重要となることに注意したい。(式 (4.2) 参照) 境界値を書いたのは補助教材に従ったまでである。

8.2 平均偏差と四分位範囲

(3) 資料の代表値としては様々なものを学習しているが, 散らばりについては範囲だけしか出てこない。統計的には分散や標準偏差が重要と思われるが, 2 次式や平方根で定義されていて, 中学 1 年生の教材としては無理と思われる。

散らばりを表す量として, ここでは平均偏差と四分位範囲について簡単に記す。

(4) 各測定値から平均値を引いたものを偏差とよぶ。平均偏差は偏差の絶対値を平均したものである。

$$(\text{平均偏差}) = \frac{\{|\,(\text{測定値}) - (\text{平均})\,|\}\text{ の測定値に関する和}}{(\text{測定値の総数})} \tag{8.1}$$

と表される。ここで $|a|$ は数 a の絶対値を表す。

度数分布表に関しては (4.1) と (4.2) の関係と同様に考えると (8.1) から

$$(\text{平均偏差}) = \frac{\{|\,(\text{階級値}) - (\text{平均})\,| \times (\text{度数})\,\}\text{ の和}}{(\text{測定値の総数})} \tag{8.2}$$

として与えられる。(8.2) の「和」は階級値に関する和とする。

平均偏差は常に正の値となる。数直線上で考えると, 平均偏差 (8.1), (8.2) は, 各測定値（あるいは各階級値）から平均値までの距離の平均を与えているので, 測定値（あるいは階級値）の散らばり具合を表す量となる。これが小さければ, 測定値は平均値の周りに密集しているし, 大きければ, 測定値は平均値から離れた点に多くあることになる。

(5) 四分位範囲は粗く言うと, 測定値の半数がその区間にあるような中央の区間の巾を求めたものである。測定値を小さい方から順に並び替えて, 全体の 1/4, 3/4 にあたる測定値をそれぞれ Q_1, Q_3 とおく。ちょうど 1/4, 3/4 にあたる測定値がなければ, その両側の隣りあった 2 数の平均を Q_1 や Q_3 にとる。

$$(\text{四分位範囲}) = Q_3 - Q_1$$

として与える。

四分位範囲も常に正となり, 範囲と同じく測定値の散らばり具合を表す量となる。ただし, 平均値とは関係のない量となる点で平均偏差とは異なる。

例 8.2 例 3.1 で与えた表 3.1 の平均偏差と四分位範囲を求めなさい。

表 3.1 を小さい順に並べたものが表 4.1 であった。4.1 節 (3) により平均値は 8.0 だったから表 4.1 の測定値から 8.0 を引いて次の表を得る。

表 **8.1**

−6	−5	−5	−4	−3	−3	−2	−2
−2	−1	−1	0	0	1	1	1
1	1	1	1	2	2	2	3
3	4	4	7				

これらの値の絶対値を平均する。平均偏差を d で表すと

$$\begin{aligned} d &= \frac{1}{28}(6+5+5+4+3+\cdots+4+7) \\ &= \frac{68}{28} \fallingdotseq 2.4 \end{aligned}$$

となる。

測定値の個数は 28 であり 4 分割すると 7 である。よって全体の 1/4, 3/4 にあたる番号はなく，表 4.1 の 7, 8 番目の平均と, 21,22 番目の平均をとって，$Q_1 = 6$, $Q_3 = 10$ となる。四分位範囲は $Q = 10 - 6 = 4$ である。

8.3 資料の相対度数と確率

(6) 表 4.1 には, 2, 4, 15 冊を借りた生徒が 1 人, 3, 5, 7, 8, 11, 12 冊が 2 人, 6, 10 冊が 3 人, 9 冊が 7 人いる。関数 p を

$$p\,(\text{借りた本の冊数}) = \frac{\text{その冊数を借りた生徒の数}}{28}$$

とすると

$$\begin{aligned} &p(0) = p(1) = p(13) = p(14) = 0 \\ &p(2) = p(4) = p(15) = \frac{1}{28} \\ &p(3) = p(5) = p(7) = p(8) = p(11) = p(12) = \frac{2}{28} \\ &p(6) = p(10) = \frac{3}{28} \qquad p(9) = \frac{7}{28} \end{aligned}$$

と表される。

このクラスの生徒を任意に 1 人とりだすとき, その人が図書館で借りた本の冊数を確率変数 X で表せば $X = k$ となる確率が

$$P(\,X = k\,) = p(k), \qquad k = 0, 1, 2, \cdots, 15$$

として定まることになる。これより

$$\begin{aligned} P(\,2 \leqq X \leqq 4\,) &= p(2) + p(3) + p(4) \\ &= \frac{1+2+1}{28} = \frac{4}{28} \\ P(\,5 \leqq X \leqq 7\,) &= p(5) + p(6) + p(7) \\ &= \frac{2+3+2}{28} = \frac{7}{28} \\ P(\,8 \leqq X \leqq 10\,) &= p(8) + p(9) + p(10) \\ &= \frac{2+7+3}{28} = \frac{12}{28} \\ P(\,11 \leqq X \leqq 13\,) &= p(11) + p(12) + p(13) \\ &= \frac{2+2+0}{28} = \frac{4}{28} \\ P(\,14 \leqq X \leqq 16\,) &= p(14) + p(15) \\ &= \frac{0+1}{28} = \frac{1}{28} \end{aligned}$$

となる。右辺の 28 を分母とする数値は相対分布表を与えている。

つまり相対度数分布は階級値の確率分布を与える。また, 相対度数表を導くことは確率の和を計算していることになる。相対度数分布と第 2 学年で学習する確率は深い関係にある。

8.4 簡便計算法

(7) 簡便計算法は平均や分散を計算する手間を軽減するための工夫である。

[注 分散は平均偏差と似ていて, 絶対値の代わりに 2 乗をとる。

$$(\text{分散}) = \frac{\{((\text{測定値}) - (\text{平均}))^2\}\text{の測定値に関する和}}{(\text{測定値の総数})}$$

と定義される。これを変形すると

$$(\text{分散}) = \frac{\{(\text{測定値})^2\}\text{の測定値に関する和}}{(\text{測定値の総数})} - \{(\text{平均値})^2\} \tag{8.3}$$

が成り立つ。]

数 a と数 $b \neq 0$ をうまくとって, すべての測定値について

$$(\text{加工値}) = \frac{(\text{測定値}) - a}{b} \tag{8.4}$$

とおき, 加工値がなるべく計算しやすい値となるようにする。(「加工」という言葉は筆者がここだけで用いたものである。) (8.4) の a を仮り平均とよぶこともある。加工値によってできる資料を加工資料とよぶことにする。このとき

$$\begin{cases} (\text{資料の平均値}) = a + b \times (\text{加工資料の平均値}) \\ (\text{資料の分散}) = b^2 \times (\text{加工資料の分散}) \\ (\text{資料の平均偏差}) = |b| \times (\text{加工資料の平均偏差}) \end{cases}$$

が成り立つ。これより, 加工資料の平均値, 分散, 平均偏差を求めれば, 与えられた資料のそれぞれの値が簡単に求まる。

例 8.3 次の資料は学生 8 人の血清総コレステロールを調べたものである。平均値, 分散, 平均偏差を求めなさい。(単位 $mg/d\ell$)

表 8.2

214	196	210	182	168	224	180	206

測定値が 200 前後の数で偶数になっているので (8.4) で $a = 200$, $b = 2$ とおいて次の表を作成する。

表 8.3

資料	加工値	2 乗	| 偏差 |
214	7	49	8.25
196	−2	4	0.75
210	5	25	6.25
182	−9	81	7.75
168	−16	256	14.75
224	12	144	13.25
180	−10	100	8.75
206	3	9	4.25
合計	−10	668	64.00

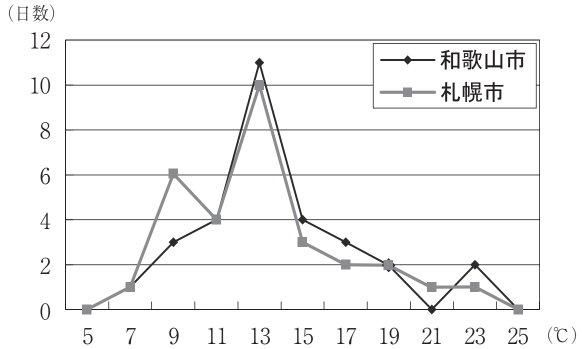
これより加工資料の平均値は $-10/8 = -1.25$ であり, 分散は (8.3) を用いて $668/8 - 1.25^2 \fallingdotseq 81.94$ となる。| 偏差 | は | (加工値) − (− 1.25) | によって計算され, 加工平均偏差は $64/8 = 8.0$ となる。

よって上記公式により, 与えられた資料の平均値は 197.5, 分散は 327.8, 平均偏差は 16.0 となる。

簡便計算法は統計ではよく知られている。例えば [7], [8] を参照されたい。

第 **7** 節の解答例

(i) 平均値 和歌山市 13.9 ℃, 札幌市 13.1 ℃, 範囲 和歌山市 16.4 ℃, 札幌市 16.9 ℃

(ii)


(iii) (1) 2 つの市の 30 日間の最高気温は 6.2 ℃から 23.4 ℃までにあり, 比較的温暖である。

(2) 2 つの度数分布多角形は, 一部を除いてよく似ている。

(3) 札幌市の方が 11 ℃ に満たない日が 3 日多く平均も 0.8 ℃低いので, 和歌山市よりも少し寒い印象を与えている。

(iv) (1) 2 つの市とも日による変化が大きい。札幌市の方がその傾向が強い。

(2) 和歌山市の方は開花が近づくにつれて暖かくなる傾向がある。

(3) 札幌市の方は開花日の前にかなり冷え込んだ。

(v) (言葉にすると上記のことでも様々な表現になりうる。)

参考文献

[1] 岡本和夫他著, 未来へひろがる数学 [1] 平成 21 年度用 補助教材 (啓林館)
[2] 重松敬一他著, 中学数学 1 平成 21 年度用 補助教材 (日本文教出版)
[3] 池田敏和他著, 中学数学 1 平成 21 年度用 補助教材 (学校図書)
[4] 文部科学省, 中学校学習指導要領 平成 21 年 3 月
[5] 文部科学省, 中学校学習指導要領解説 平成 21 年 7 月
[6] 文部科学省, 小学校学習指導要領 平成 21 年 7 月
[7] 稲垣宣夫, 山根芳和, 吉田光雄, 統計学入門 (裳華房) 1992
[8] 北川敏男, 稲葉三男, 統計学通論 (第 2 版) (共立出版) 1960